



# Metal-histidine-glutamate as a regulator of enzymatic cycles: a case study of carbonic anhydrase.

Gilles Frison, Gilles Ohanessian

## ► To cite this version:

Gilles Frison, Gilles Ohanessian. Metal-histidine-glutamate as a regulator of enzymatic cycles: a case study of carbonic anhydrase.. *Physical Chemistry Chemical Physics*, 2009, 11 (2), pp.374-383. 10.1039/b812916a . hal-00467064

**HAL Id: hal-00467064**

**<https://hal.science/hal-00467064>**

Submitted on 25 Mar 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Metal-histidine-glutamate as a regulator of enzymatic cycles: a case study of carbonic anhydrase†

Gilles Frison\* and Gilles Ohanessian

Received 30th July 2008, Accepted 24th September 2008

First published as an Advance Article on the web 5th November 2008

DOI: 10.1039/b812916a

Histidine is a very common metal ligand in metalloenzymes. Besides being an efficient Lewis base, its electronic properties are essential to shape the metal ability to catalyze the reaction. Here we show that histidine's properties can be tuned, in turn, by an easy proton transfer to a nearby glutamate. We study this situation in Human Carbonic Anhydrase II (HCA II) in which one of the three histidines bound to zinc (His119) interacts also with a glutamate residue (Glu117). Proton transfer from His119 to Glu117 has been hypothesized in the past, however realistic modeling is performed here for the first time. We show that the carboxylate group of Glu117 behaves only as a hydrogen bond acceptor in the hydroxy form of HCA II. On the other hand, our results suggest that Glu117 could exist either as a hydrogen bond acceptor or as a proton acceptor in the aqua form of HCA II, the two isomers having almost the same thermodynamic stability. We propose that this proton shift may be used by the enzyme to facilitate the final displacement of bicarbonate by water.

## Introduction

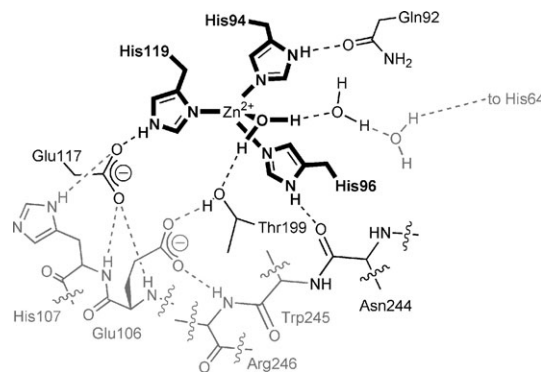
Carbonic anhydrase (CA), the first enzyme known to require zinc as a cofactor, catalyzes the reversible hydration of carbon dioxide.<sup>1</sup> This fundamental reaction, involved in many physiological processes such as respiration, photosynthesis or acid–base balance of all living organisms, has put CA in a prominent place in structure–function studies of zinc enzymes. Of the eleven active isozymes known in the  $\alpha$ -family of CA, most of the chemical, structural and computational studies have been done on the human CA II (HCA II) enzyme.<sup>2</sup>

HCA II is one of the most efficient biological catalysts known with a turnover rate of  $10^6 \text{ sec}^{-1}$  at pH = 9 and 25 °C, and a three main step mechanism (proton release from Zn-bound water to form a Zn-bound hydroxide; nucleophilic attack by the Zn–OH unit on the carbon atom of  $\text{CO}_2$ ; replacement of the Zn-bound bicarbonate by an external water molecule) has been established by experimental and theoretical investigations.<sup>2,3</sup> By 1992, the crystal structure of HCA II has been refined to 1.54 Å resolution.<sup>4</sup> Zinc is coordinated to the nitrogen atoms of His94, His96 and His119 and the oxygen atom of a water molecule to form an approximately tetrahedral complex (Fig. 1). Due to the high Lewis-acidity of the metal, the  $\text{p}K_{\text{a}}$  of the Zn-bound water is drastically reduced to near 7, making its deprotonation possible.

The  $\text{Zn}(\text{His})_3(\text{H}_2\text{O})$  active site is incorporated into a hydrogen bond network that includes: (i) the “gatekeeper” Thr199

and Glu106, hydrogen-bonded to the Zn-bound  $\text{H}_2\text{O}$ ; (ii) Gln92 sidechain, Asn244 peptide linkage, and Glu117 sidechain hydrogen-bonded to the second nitrogen of the imidazole moiety of His94, His96 and His119 respectively; (iii) a chain of water molecules connecting the active site to the proton acceptor His64 (Fig. 1).

X-Ray and NMR structures of metalloenzymes<sup>5</sup> show the insertion of their active site into a complex hydrogen bond network.<sup>6</sup> In most cases however, the hydrogen atom of the H-bond is not located and thus its directionality is not known with certainty. Thus the hydrogen atoms' location on the protein residues is based on their acidity. Among ionizable groups in proteins, deprotonation of imidazole, which produces an imidazolate anion, is not possible in bulk water<sup>7</sup> but has been proposed in case of metal coordination.<sup>8,9</sup> Several experimental studies based on NMR,<sup>10</sup> resonance Raman<sup>11</sup> or Fourier transform infrared<sup>12–14</sup> spectroscopy suggest the formation of iron-imidazolate moiety in heme



**Fig. 1** Amino acid residues and hydrogen bonds at the active center of HCA II. Bold: first shell ligands, black: second shell ligands, grey: third and further shell ligands.

Laboratoire des Mécanismes Réactionnels, Département de Chimie, Ecole Polytechnique and CNRS, 91128 Palaiseau Cedex, France.  
E-mail: gilles.frison@polytechnique.org; Fax: +33 1 69 33 48 03;  
Tel: +33 1 69 33 48 34

† Electronic supplementary information (ESI) available: Table of  $\text{Zn}-\text{C}(\text{Carboxylate})$  interatomic distances. Cartesian coordinates, single point energies, gas-phase and solvated free energies of the stationary points. See DOI: 10.1039/b812916a

and non heme proteins. Notably, the presence of the Fe-His-Asp/Glu triad, associated with the low  $pK_a$  value of the metal-bound histidine,<sup>15</sup> is consistent with a deprotonated His ligand. In heme proteins, this is associated with the charge relay mechanism where the hydrogen bond network around the iron-bound histidine tunes the charge donated by the His to the Fe, and stabilize oxidation states greater than Fe(III).<sup>16–18</sup>

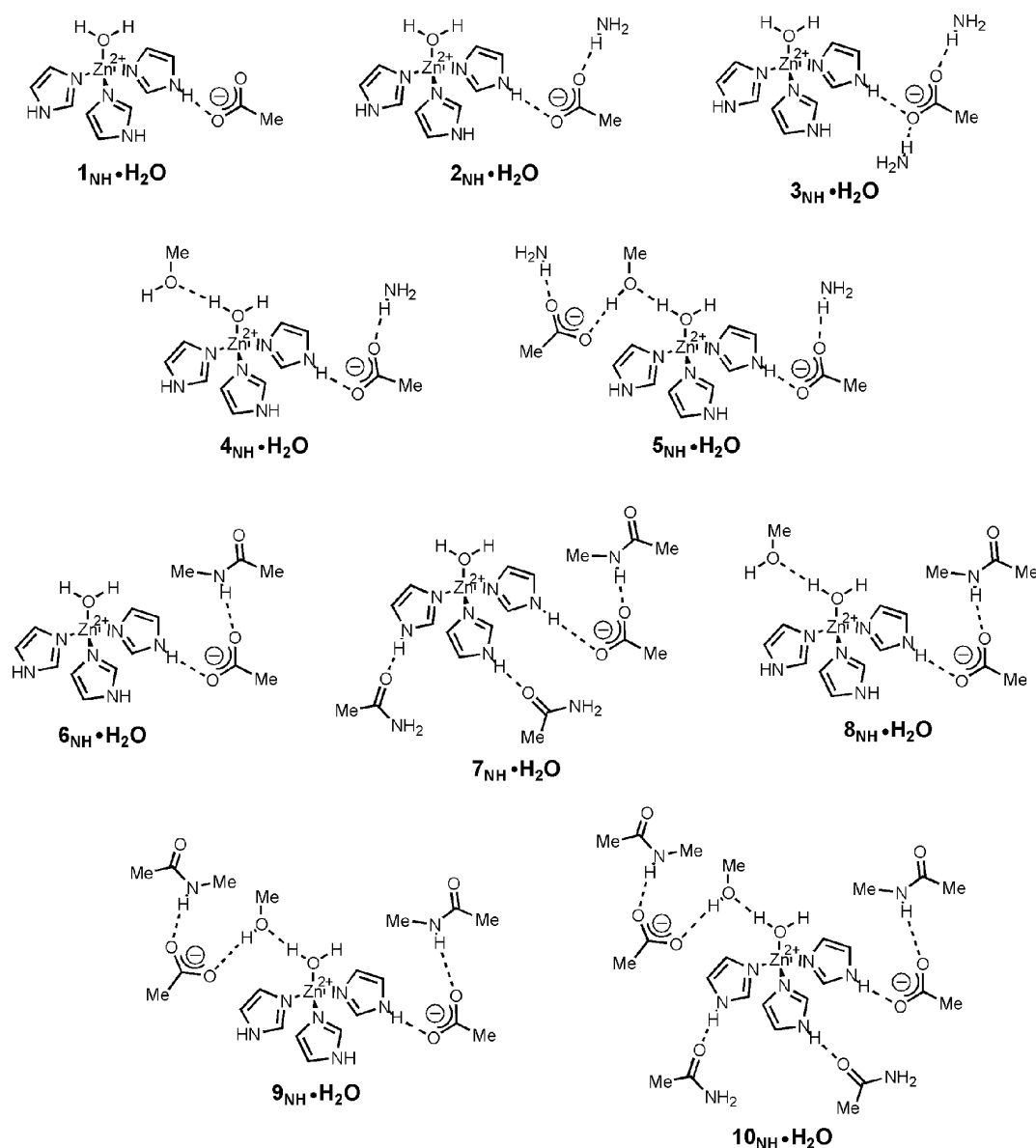
Previous studies have suggested that a Zn-bound His could also exist in a deprotonated form when hydrogen-bonded to Asp or Glu.<sup>8,19–26</sup> It has been shown that this propensity is dependent on the nature of the other zinc ligands and the solvent accessibility of the Zn-binding site.<sup>24</sup> For complex **1**·H<sub>2</sub>O (Scheme 1), a simple model of the HCA II active site, it has been calculated for both gas phase and solvated cases, that the proton is preferentially located on the acetate rather than on the imidazole. This result is however conflicting with the “universal” description of the HCA

II active site with three neutral imidazoles linked to the metal. In this work, we examine the question of a possible shuttling between hydrogen-bonded atoms using models that are more realistic than in previous studies, with the help of high level computational methods. Several approaches have been taken during this work, from small to medium sized models optimized at the HF or B3LYP levels of calculation to QM/QM' modeling of a large fraction of the enzyme.

## Results

### 1. Flexible models of the HCA II active site

Several models of the HCA II active site containing up to 93 atoms have been constructed (Scheme 1). First, we consider two different first shell ligand sets, three imidazoles (Im) and one water molecule (noted as ·H<sub>2</sub>O), and three imidazoles and one hydroxy group (noted as ·OH<sup>−</sup>). This allows us to



**Scheme 1** Flexible models of the HCA II active site examined in this work. Only ·H<sub>2</sub>O compounds and NH isomers are drawn.

examine two successive minima of the active site along the reaction process, corresponding to the extreme (neutral water and ionic hydroxy) electronic influence of the oxygen ligand on the metal. Second, on the basis of the X-ray structure (PDB Code 2CBA),<sup>4</sup> we have incorporated the effect of the hydrogen bond network all around the active site, in a stepwise manner. This includes groups modelling second shell (Gln92, Glu117, Thr199 and Asn244) and third and further shell (Glu106, His107 and Arg246) ligands (Fig. 1). It should be noted that these neighboring groups are not specific of the HCA II isozyme<sup>4</sup> and are shared by all the mammalian carbonic anhydrase isozymes for which an X-ray crystallographic structure has been obtained.<sup>27</sup> Third, two isomers have been examined, one containing a  $\text{Zn}^{2+}\text{-ImH}^0\cdots\text{MeCOO}^-$  unit (NH isomer, noted as NH suffix), the other containing a  $\text{Zn}^{2+}\text{-Im}^-\cdots\text{MeCOOH}^0$  unit (OH isomer, noted as OH suffix).

Depending upon the model, NH peptidic groups of Glu106, His107 and Arg246 have been included in order to add hydrogen bonds to the Glu106 and Glu117 carboxylates and thus properly describe their electronic influence in the enzyme. They have been modelled either by  $\text{NH}_3$  (in **2**–**5**) or by *N*-methyl-acetamide (in **6**–**10**). Compounds **4**, **5** and **8**–**10** involve a hydrogen bond between the Zn-bound oxygen ligand and a methanol modelling the Thr199 residue. The “gate-keeper” has been included in our models as it has been shown to govern the rearrangement of the zinc-coordinated  $\text{HCO}_3^-$  ion during the catalytic  $\text{CO}_2$  hydration cycle of HCA II.<sup>28</sup> Compound **3** has been built in order to model the experimentally observed hydrogen bonds on both oxygen atoms of the acetate/acetic acid moiety (Ace) of Glu117.

Geometry optimization of these models has been conducted without any restriction, thus allowing to estimate entropic and thermal effects from a harmonic analysis of the normal-mode frequencies of the system. In a number of cases, geometry optimization of these flexible models leads to significant deviation of the conformation compared to the X-ray structure because others type of hydrogen bonding interactions are locally more favorable. In all such instances, the models have been excluded and the results will not be described below.

For several cases (**1**· $\text{H}_2\text{O}$ , **2**· $\text{H}_2\text{O}$  and **4**· $\text{H}_2\text{O}$ ), the NH isomer could not be located as a minimum, the optimization process leading back to the OH isomer (see Methods). The same result is obtained for **2**· $\text{OH}^-$  but only at the B3LYP level. In such cases, the NH isomer geometry has been obtained by constraining the NH bond length at 1.06 Å.

The effects of the various shell ligands around zinc on the gas-phase energy difference between OH and NH isomers are given in Table 1.

**1** is the simplest complex in our study modeling the Zn-His119-Glu117 triad present in the HCA II active site. As in a previous study on restrained optimized structures of **1**· $\text{H}_2\text{O}$  and **1**· $\text{OH}^-$ ,<sup>24</sup> we find that the OH isomer is much lower in energy than the NH isomer in both cases. Furthermore, as already noted, **1**<sub>NH</sub>· $\text{H}_2\text{O}$  could not be located on the potential energy surface and has been optimized by constraining the NH bond length at 1.06 Å. Moving from the · $\text{H}_2\text{O}$  to the · $\text{OH}^-$  complex induces a decrease (from 55.1 to 19.5 kJ mol<sup>−1</sup>) of the relative stability of the OH isomer compared to the NH. The  $\text{Im}^-\cdots\text{MeCOOH}^0$  moiety is more stabilized by a high Lewis

**Table 1** Relative gas-phase enthalpy  $\Delta H$  and free energy  $\Delta G^1$  of the NH isomer compared to the OH one (in kJ mol<sup>−1</sup> at 298 K) at the B3LYP/BS2//HF/BS1 level<sup>a</sup>

| Species               | · $\text{H}_2\text{O}$ compounds |                           | · $\text{OH}^-$ compounds |                           |
|-----------------------|----------------------------------|---------------------------|---------------------------|---------------------------|
|                       | $\Delta H$                       | $\Delta G^{1b}$           | $\Delta H$                | $\Delta G^{1b}$           |
| <b>1</b>              | 55.1 (59.4 <sup>c</sup> )        | 67.3 (71.6 <sup>c</sup> ) | 19.5 (24.3 <sup>c</sup> ) | 22.9 (27.7 <sup>c</sup> ) |
| <b>2<sup>d</sup></b>  | 47.6 (50.9 <sup>c</sup> )        | 50.8 (54.0 <sup>c</sup> ) | 12.8 (16.3 <sup>c</sup> ) | 17.9 (21.4 <sup>c</sup> ) |
| <b>3</b>              | — <sup>e</sup>                   | — <sup>e</sup>            | −2.5 (9.5 <sup>c</sup> )  | 6.7 (18.6 <sup>c</sup> )  |
| <b>4</b>              | 40.1 (43.4 <sup>c</sup> )        | 56.2 (59.5 <sup>c</sup> ) | — <sup>e</sup>            | — <sup>e</sup>            |
| <b>5</b>              | 20.0                             | 23.3                      | — <sup>e</sup>            | — <sup>e</sup>            |
| <b>6</b>              | 23.9 (26.9 <sup>c</sup> )        | 29.9 (32.9 <sup>c</sup> ) | 10.4 (13.0 <sup>c</sup> ) | 11.7 (14.3 <sup>c</sup> ) |
| <b>7</b>              | 18.0                             | 28.8                      | 3.2                       | 5.6                       |
| <b>8</b>              | 25.2                             | 32.3                      | — <sup>e</sup>            | — <sup>e</sup>            |
| <b>9</b>              | 13.3                             | 18.3                      | — <sup>e</sup>            | — <sup>e</sup>            |
| <b>10</b>             | −10.9                            | −0.5                      | — <sup>e</sup>            | — <sup>e</sup>            |
| <b>11<sup>f</sup></b> | 19.5 <sup>g</sup>                | 29.8 <sup>g</sup>         |                           |                           |
| <b>12<sup>h</sup></b> | 42.6 <sup>g</sup>                | 52.9 <sup>g</sup>         |                           |                           |

<sup>a</sup> A negative value indicates that the NH isomer is more stable than the OH isomer. <sup>b</sup> Gas-phase free energy correction computed at the HF/BS1 level. <sup>c</sup> MP2/BS2//HF/BS1 value. <sup>d</sup>  $\Delta H = 47.5$  (12.3 resp.) kJ mol<sup>−1</sup> and  $\Delta G^1 = 56.1$  (23.4 resp.) kJ mol<sup>−1</sup> for **2**· $\text{H}_2\text{O}$  (**2**· $\text{OH}^-$ , respectively) at the B3LYP/BS2//B3LYP/BS1 level. <sup>e</sup> No relevant structure obtained. <sup>f</sup>  $\Delta H = 22.2$  kJ mol<sup>−1</sup> and  $\Delta G^1 = 32.6$  kJ mol<sup>−1</sup> at the B3LYP/BS2//B3LYP/BS1 level. <sup>g</sup> Enthalpies and gas phase free energies have been derived from thermal and entropic contributions computed for **10**. <sup>h</sup> Relative energy computed only for the inner shell of **12**.

acidity of Zn, and thus by a low charge transfer from first-shell ligands around the metal.<sup>24</sup>

Compared to **1**, all others complexes included in this part show smaller energy differences between the OH and NH isomers. In all cases except **3**· $\text{OH}^-$  and **10**· $\text{H}_2\text{O}$ , the OH isomers remain lower in energy than the NH isomers, however. This reduced difference is due to the inclusion of chemical groups that establish hydrogen-bonding interactions with **1**. Indeed, hydrogen bonds on the Zn-bound water or on the imidazole induce negative charge transfer from these direct metal ligands to Zn, lowering its Lewis acidity. On the other hand, hydrogen bonds on the oxygen of the acetate (or acetic acid) group decrease its basicity and thus its capability to deprotonate the imidazole ring.

The progressive inclusion of these hydrogen bonds shows the relative influence of each, as well as their cumulative effect. Adding a hydrogen bond on the second oxygen atom of the carboxylate group through  $\text{NH}_3$  (comparison between **2** and **1**) or on the Zn-bound water molecule through methanol (comparison between **4** and **2** or between **8** and **6**) has only a small effect on the relative stability of NH vs. OH isomers. Indeed, for **2**· $\text{H}_2\text{O}$  and **4**· $\text{H}_2\text{O}$ , as for **1**· $\text{H}_2\text{O}$ , the NH isomer could not be located on the potential energy surface and has been optimized by constraining the NH bond length at 1.06 Å. The effect seems to be more pronounced when *N*-methyl acetamide is used instead of  $\text{NH}_3$  (comparison between **6** and **2**). This is especially the case for **6**· $\text{H}_2\text{O}$  due to an additional hydrogen bond formed between the oxygen atom of the *N*-methyl acetamide and the Zn-bound water. Adding a hydrogen bond to the methanol, linked to the  $\text{Zn}^{2+}\text{-OH}_2$  moiety, through a negatively charged methyl acetate (comparison between **5** and **4** or between **9** and **8**) leads to a stronger stabilization of the NH isomers compared

to the OH ones. The inclusion of a hydrogen bond on the oxygen atom of the carboxylate group that is hydrogen bound to the imidazole group (comparison between **3** and **2**) is of quite the same amount. The effect of two acetamide hydrogen bonds to Zn-bound imidazoles on the relative stability of NH vs. OH isomers is small when moving from **6** to **7**. It is clearly more marked when comparing complexes **10** and **9**, due to an additional hydrogen bond formed between the acetate group of the  $\text{Zn}^{2+}$ -imidazole-acetate triad and one of the supplementary acetamides.

In all cases, inclusion of the entropy corrections leads to gas-phase free energy differences  $\Delta G^\ddagger$  which show an increase of the relative stability of the OH isomer compared to the NH isomer. The higher flexibility of the OH isomer due to weaker  $\text{Zn}^{2+}$ -Im $^-$ ...MeCOOH $^0$  hydrogen bond compared to  $\text{Zn}^{2+}$ -ImH $^0$ ...MeCOO $^-$  one, as illustrated by longer H-bond (Fig. 2,  $7_{\text{NH}}\cdot\text{H}_2\text{O}$  vs.  $7_{\text{OH}}\cdot\text{H}_2\text{O}$  and  $7_{\text{NH}}\cdot\text{OH}^-$  vs.  $7_{\text{OH}}\cdot\text{OH}^-$ ) explains this variation (see above).

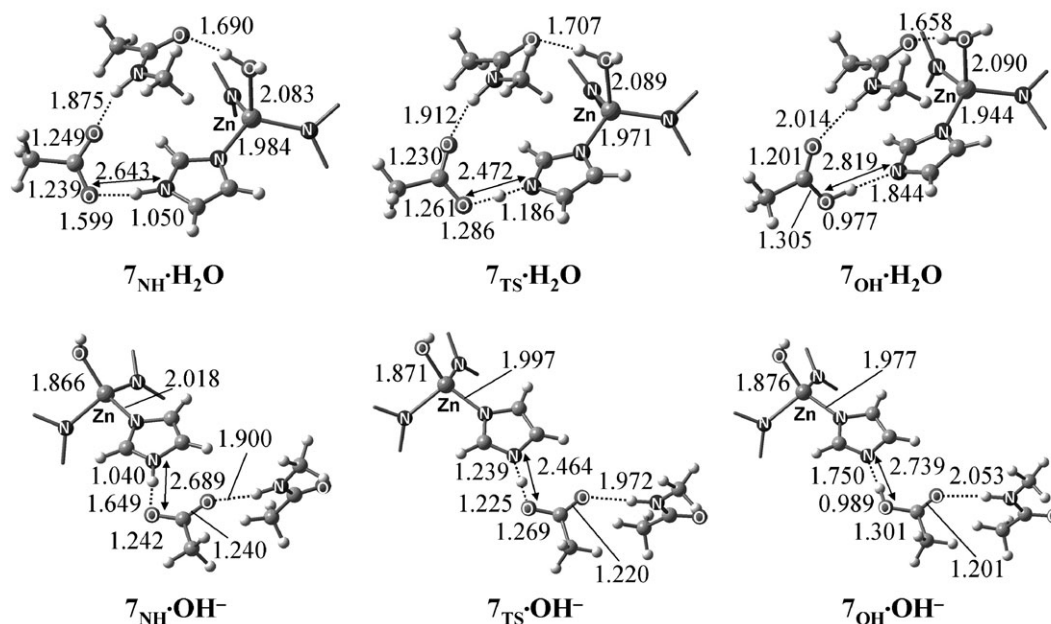
Solvation effects on the relative stability of NH and OH isomers are given in Table 2. It has been shown previously on model **1** that an increase of the dielectric constant (from 1 to 80) favors the NH isomer mainly because the outer-shell carboxylate moiety is better solvated than the outer-shell protonated COOH group.<sup>24</sup>

This effect is reproduced by our calculations, leading to a NH isomer of **1**·H $_2$ O more stable by 5 kJ mol $^{-1}$  than the OH isomer in case of water solvation. The water solvation free energy term of the NH isomer actually exceeds the OH one by more than 70 kJ mol $^{-1}$ . The same quantitative effect is observed for **2**·H $_2$ O and **4**·H $_2$ O which correspond to the smaller models. The increase of the model size leads to a reduction of the solvation effect. Indeed, the water solvent effect rises to about 32 kJ mol $^{-1}$  for **5**–**6**·H $_2$ O and **8**·H $_2$ O, to ca. 15 kJ mol $^{-1}$  for **7**·H $_2$ O and **9**·H $_2$ O, and only to 5 kJ mol $^{-1}$  for **10**·H $_2$ O. This leads to the result that while small ·H $_2$ O

models have their NH isomer as the absolute minimum in water, larger ·H $_2$ O models do not show a clear preference between OH and NH isomers (Fig. 3). It is as if increasing the reliability of the ·H $_2$ O model would conduct to two isoenergetic isomers in case of water solvation. Buried sites, characterized by a dielectric constant of 4, present the same feature with a convergence to almost two isoenergetic isomers.

The ·OH $^-$  models gives less diverse results. Indeed, irrespective of the model considered, the NH isomer is more stable than the OH isomer by ca. 10 kJ mol $^{-1}$  for buried sites and by 15–19 kJ mol $^{-1}$  in water.

The energetic similarity of the two isomers for large models led us to investigate their interconversion barrier. This has been conducted on **7**, the largest complex for which transition states between OH and NH isomers (noted as TS suffix) could be located for both the ·H $_2$ O and the ·OH $^-$  models. Fig. 2 illustrates the effect of proton transfer on the main geometrical parameters of the two minima and the transition state for both **7**·H $_2$ O and **7**·OH $^-$ . The calculated Im(N)...Ace(O) hydrogen bond lengths (between 2.643 and 2.819 Å) are in good agreement with the experimental data of His(N)...Asp/Glu(O) distances (between 2.6 and 2.9 Å) observed in crystallographic structures of zinc enzymes including a Zn-His-Asp/Glu triad.<sup>26</sup> This is clearly shorter than the constrained hydrogen bond length (around 3.3 Å) used in a previous study.<sup>24</sup> The shorter H-bond lengths are observed for the NH isomer, due to electrostatic effects. Indeed, in the NH isomer, this interaction takes place between a negatively charged acetate group and a partially positive (due to charge transfer to  $\text{Zn}^{2+}$ ) imidazole group. This electrostatic interaction almost disappears in OH isomers as the acetate group is protonated. The transition states show a net shortening (around 0.2–0.3 Å) of the Im(N)...Ace(O) distance, which should facilitate proton transfer between Im(N) and Ace(O).

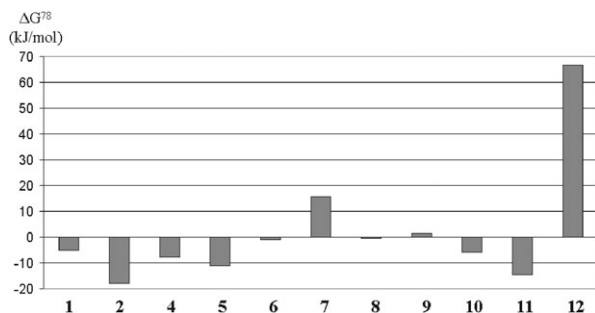


**Fig. 2** Main geometrical parameters (bond lengths in Å) of optimized structures of **7**-type species at the HF/BS1 level. The two imidazole-acetamide groups, except for Zn-bound nitrogen, have been omitted for clarity.

**Table 2** Relative free energy of the NH and OH isomers (in kJ mol<sup>-1</sup>)<sup>a</sup>

| Species | <b>·H<sub>2</sub>O</b> compounds |                             |                    | <b>·OH<sup>-</sup></b> compounds |                             |                    |
|---------|----------------------------------|-----------------------------|--------------------|----------------------------------|-----------------------------|--------------------|
|         | $\Delta G^4$                     | $\Delta G^{78}$             | $\Delta\Delta G^b$ | $\Delta G^4$                     | $\Delta G^{78}$             | $\Delta\Delta G^b$ |
| 1       | 12.8 (17.1 <sup>c</sup> )        | -5.0 (-0.6 <sup>c</sup> )   | -72.3              | -8.9 (-4.1 <sup>c</sup> )        | -18.9 (-14.1 <sup>c</sup> ) | -41.8              |
| 2       | -1.8 (1.5 <sup>c</sup> )         | -18.0 (-14.7 <sup>c</sup> ) | -68.8              | -7.4 (-3.9 <sup>c</sup> )        | -15.2 (-11.7 <sup>c</sup> ) | -33.1              |
| 3       | — <sup>d</sup>                   | — <sup>d</sup>              | — <sup>d</sup>     | -11.0 (1.0 <sup>c</sup> )        | -15.5 (-3.5 <sup>c</sup> )  | -22.1              |
| 4       | 7.3 (10.6 <sup>c</sup> )         | -7.7 (-4.4 <sup>c</sup> )   | -63.9              | — <sup>d</sup>                   | — <sup>d</sup>              | — <sup>d</sup>     |
| 5       | -2.8                             | -10.9                       | -34.2              | — <sup>d</sup>                   | — <sup>d</sup>              | — <sup>d</sup>     |
| 6       | 7.9 (10.9 <sup>c</sup> )         | -1.0 (1.9 <sup>c</sup> )    | -31.0              | -10.0 (-7.5 <sup>c</sup> )       | -17.3 (-14.7 <sup>c</sup> ) | -29.0              |
| 7       | 20.3                             | 15.7                        | -13.1              | -11.5                            | -17.6                       | -23.2              |
| 8       | 8.7                              | -0.4                        | -32.7              | — <sup>d</sup>                   | — <sup>d</sup>              | — <sup>d</sup>     |
| 9       | 7.1                              | 1.7                         | -16.6              | — <sup>d</sup>                   | — <sup>d</sup>              | — <sup>d</sup>     |
| 10      | -4.9                             | -5.9                        | -5.3               | — <sup>d</sup>                   | — <sup>d</sup>              | — <sup>d</sup>     |
| 11      | -2.5                             | -14.5                       | -44.3              | — <sup>d</sup>                   | — <sup>d</sup>              | — <sup>d</sup>     |
| 12      | 177.6 <sup>e</sup>               | 66.7 <sup>f</sup>           | +13.8              |                                  |                             |                    |

<sup>a</sup> Gas-phase free energy computed at 298 K at the B3LYP/BS2//HF/BS1 level, plus the solvation free energy correction computed at the CPCM-HF/BS1 + level with a dielectric constant of 4 and 78.39 for  $\Delta G^4$  and  $\Delta G^{78}$ , respectively. A negative value indicates that the NH isomer is more stable than the OH isomer. <sup>b</sup>  $\Delta\Delta G = \Delta\Delta G^{78} - \Delta\Delta G^4$ . <sup>c</sup> Gas-phase free energy computed at the MP2/BS2//HF/BS1 level. <sup>d</sup> No relevant structure obtained. <sup>e</sup> Computed at the B3LYP level with BS2 and STO-3G used for atoms belonging to the inner and outer layers, respectively. A value of 195.6 kJ mol<sup>-1</sup> is obtained at the HF level with the BS2 and 3-21G composite basis set. <sup>f</sup> Solvation free energy correction computed at the CPCM-PM3MM level.

**Fig. 3** Free energy difference between the OH and the NH isomers of the **·H<sub>2</sub>O** models. A negative value indicates that the NH isomer is more stable than the OH isomer.

This transfer is calculated to have a small activation barrier at the optimisation level, with  $\Delta E$  values of 5.6 and 13.0 kJ mol<sup>-1</sup> respectively for **7·H<sub>2</sub>O** and **7·OH<sup>-</sup>** (Table 3). Furthermore, including correlation effects, thermal enthalpic, entropic and solvation corrections leads to the vanishing of any barrier for proton transfer, the **7<sub>TS</sub>** transition state being calculated more stable than the NH isomer (in the case of **7·H<sub>2</sub>O**) or even more stable than both NH and OH isomers (in the case of **7·OH<sup>-</sup>**). Even though the proton could be located near the

N(Im) or the O(Ace) atom, the potential energy surface around these positions seems to be very flat, thus allowing a fine tuning of the charge donated by the His to the Zn at a low energetic cost.

## 2. Constrained model of the HCA II active site

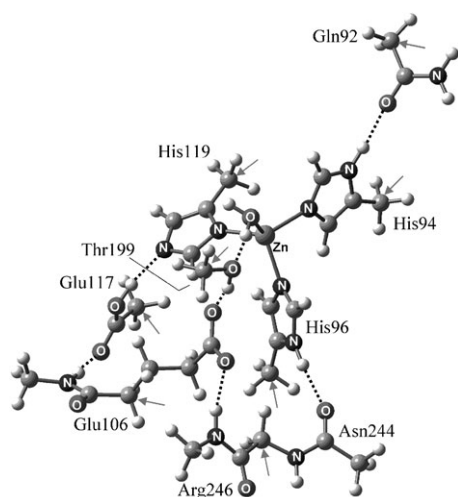
Flexible models do not introduce bias in the optimization process and permit to evaluate the thermal and entropic contributions necessary for the calculation of enthalpies and free energies. However, they lead in some cases to strong distortion of the model geometries compared to the X-ray structure of the active site embedded in the protein network. A simple way around this problem is to use experimental data to set up constraints on the geometry. We have built a new model of 101 atoms, labelled **11**, mostly similar to our larger flexible model **10** (Fig. 4). **11·H<sub>2</sub>O** has been studied in its OH and NH isomer forms by fixing eight alkyl carbons, located in the periphery of the model, in their relative position in the X-ray structure (PDB code 2CBA). The rest of the structure has been fully optimised at the HF/BS1 and B3LYP/BS1 levels.

At the B3LYP/BS2//HF/BS1 level, we find that the **11<sub>OH</sub>·H<sub>2</sub>O** isomer is located 19.5 kJ mol<sup>-1</sup> lower in enthalpy than the **11<sub>NH</sub>·H<sub>2</sub>O** isomer (Table 1). The same result is obtained at the B3LYP/BS2//B3LYP/BS1 level, with an enthalpy difference

**Table 3** Relative energy of the isomers of **7** (in kJ mol<sup>-1</sup>)

| Species          | <b>·H<sub>2</sub>O</b> compounds |                       |                       | <b>·OH<sup>-</sup></b> compounds |                       |                       |
|------------------|----------------------------------|-----------------------|-----------------------|----------------------------------|-----------------------|-----------------------|
|                  | <b>7<sub>NH</sub></b>            | <b>7<sub>TS</sub></b> | <b>7<sub>OH</sub></b> | <b>7<sub>NH</sub></b>            | <b>7<sub>TS</sub></b> | <b>7<sub>OH</sub></b> |
| $\Delta E^a$     | 0.0                              | + 5.6                 | -24.8                 | 0.0                              | + 13.0                | -8.1                  |
| $\Delta E^b$     | 0.0                              | -10.7                 | -18.6                 | 0.0                              | -4.5                  | -2.3                  |
| $\Delta H^b$     | 0.0                              | -23.7                 | -18.0                 | 0.0                              | -19.0                 | -3.2                  |
| $\Delta G^{1b}$  | 0.0                              | -25.0                 | -28.8                 | 0.0                              | -16.6                 | -5.6                  |
| $\Delta G^{4c}$  | 0.0                              | -15.9                 | -20.3                 | 0.0                              | -5.0                  | + 11.5                |
| $\Delta G^{78c}$ | 0.0                              | -12.5                 | -15.7                 | 0.0                              | -0.4                  | + 17.6                |

<sup>a</sup> Relative energy at the HF/BS1 level. <sup>b</sup> Relative energy, enthalpy and gas-phase free energy at 298 K at the B3LYP/BS2//HF/BS1 level. <sup>c</sup> Relative free energy with solvation correction computed at the CPCM-HF/BS1 + level with a dielectric constant of 4 and 78.39 for  $\Delta G^4$  and  $\Delta G^{78}$ , respectively.



**Fig. 4** Optimized structure of  $11_{\text{OH}}\cdot\text{H}_2\text{O}$  at the HF/BS1 level. Arrows indicate the eight carbons fixed in their X-ray structure relative positions.

of  $22.2 \text{ kJ mol}^{-1}$ . It should be noted that, as observed for  $2\cdot\text{OH}^-$ , and contrary to the HF/BS1 level, the geometry optimization of  $11_{\text{NH}}\cdot\text{H}_2\text{O}$  at the B3LYP/BS1 level leads back to the  $11_{\text{OH}}\cdot\text{H}_2\text{O}$  isomer, thus requiring to fix the N–H bond length. Inclusion of entropic effects leads to free energy difference of  $29.8 \text{ kJ mol}^{-1}$ .

Compared to **10**, these results indicate that freezing the geometry partially in order to preserve the main features of the X-ray structure favors the OH isomer by *ca.*  $30 \text{ kJ mol}^{-1}$  in the gas phase. This is due to the fact that the carboxylate group in the NH isomer, modeling Glu117, could not bind the histidine group as strongly as in **10**, due to the imposed constraint.

This trend is reversed after inclusion of solvent effects (Table 2). Indeed, the constraint leads to a carboxylate group that is more exposed to the solvent, thus favoring by  $2.5 \text{ kJ mol}^{-1}$  ( $14.5 \text{ kJ mol}^{-1}$ , respectively) the NH isomer with a dielectric constant equal to 4 (78, respectively). This confirms the result obtained with the flexible models indicating that large  $\cdot\text{H}_2\text{O}$  models do not show a clear preference between OH and NH isomers, particularly for a buried site.

### 3. QM/QM' model of the HCA II active site

In order to better describe the effect of the anisotropic environment of the enzyme around its active site, we have built a model of 977 atoms, labeled **12**, including residues and water molecules possessing at least one atom within  $10 \text{ \AA}$  of the zinc atom. Model **12** has been divided into two layers, the inner layer corresponding to model **11** whereas the remainder 876 atoms belong to the outer layer. Partial geometry optimization of  $12_{\text{OH}}\cdot\text{H}_2\text{O}$  and  $12_{\text{NH}}\cdot\text{H}_2\text{O}$  has been achieved at the ONIOM(HF/BS1:PM3MM) level by fixing the relative positions of all  $\text{C}_\alpha$  carbons of the outer layer. Final energy calculation at the B3LYP/BS2 level only for the inner shell leads to results in line with those obtained with model **11**. Indeed, the free energy of the inner layer of  $12_{\text{OH}}\cdot\text{H}_2\text{O}$  is  $52.9 \text{ kJ mol}^{-1}$  lower than that of the inner layer of  $12_{\text{NH}}\cdot\text{H}_2\text{O}$ . Inclusion of the effect of the surrounding protein on the inner shell atoms, evaluated for the previous systems with a

dielectric constant of 4 at the CPCM method, has been obtained here by a single-point calculation at the B3LYP level with a composite basis set: BS2 and STO-3G were used for the inner and outer layer atoms, respectively. Contrary to the effect of a polarizable continuum, the anisotropic environment of the enzyme greatly increases the relative stability of the OH isomer, with a preference of  $177.6 \text{ kJ mol}^{-1}$  (Table 2). Modifying the basis set of the outer layer atoms to 3–21G, evaluated at the HF level, still amplified this trend by almost  $20 \text{ kJ mol}^{-1}$ . Lastly, as for smaller models **1–11**, water solvation stabilizes more strongly the NH than the OH isomer. This effect however does not reverse the relative stability of the isomers, leading to the global minimum corresponding to  $12_{\text{OH}}\cdot\text{H}_2\text{O}$  (Table 2).

## Discussion

### 1. Modelling strategy and accuracy

Quantum *ab initio* and DFT chemical studies of metallo-enzymes are limited to models built from the active site of the protein. It has however been established that the protein environment plays a noticeable influence on the structure, stability and reactivity of the enzyme. In particular, several studies suggest that the second-shell ligands around the metal,<sup>29–31</sup> and more precisely the carboxylate–histidine–zinc triad,<sup>25,32</sup> play structural (orienting the first-shell ligands) and electronic (fine-tuning the  $\text{pK}_a$  and reactivity of Zn-bound water) roles. It is thus crucial to establish the limitation and reliability of the compounds used to model metalloenzymes.

The results in Tables 1 and 2 and in Fig. 3 show the influence of the environment on the structure of the HCA II active site. First, the relative energy of NH and OH isomers in the gas-phase strongly depends upon the model used. Indeed, the B3LYP/BS2//HF/BS1 level potential energy surface presents sometimes only one minimum, the OH isomer (for **1–2·H<sub>2</sub>O** and **4·H<sub>2</sub>O**), and sometimes two minima, the absolute one being either the OH isomer (for **5–9·H<sub>2</sub>O**, **11–12·H<sub>2</sub>O**, **1–2·OH<sup>−</sup>** and **6–7·OH<sup>−</sup>**) or the NH isomer (for **10·H<sub>2</sub>O** and **3·OH<sup>−</sup>**).

Furthermore, two supplementary parameters appear to play in favor of the OH isomer: (i) the use of the more reliable<sup>33</sup> MP2 level instead of B3LYP as method for the final energy computation always increases the stability of the OH isomer relative to that of the NH isomer; (ii) the use of B3LYP instead of HF as geometry optimization method slightly favors the OH isomer. Both effects would concur to a very small energy difference between the two isomers, so that it is very difficult to ascertain which is the absolute gas-phase or solvated free energy minimum for the more reliable medium-size models **10–11**. Extended model **12** gives a more conclusive result but it should be noted that the lower level of calculation used for the outer layer increases the error bar. The methods used would thus not allow to conclude firmly about which isomer is most stable but establish clearly that the energy difference between them: (i) is strongly dependent upon the model and environment used; (ii) is relatively small. The first point confirms a very recent study showing the difficulty to converge to a proper value depending both on the model used and on the computational procedure.<sup>34</sup> We believe the second point

to be a key for controlling the catalytic cycle of the enzyme (*vide infra*).

This study sets the problem of the choice of the model in order to describe the zinc binding site of HCA II in a meaningful way. The influence of indirect metal ligands on the reactivity of HCA II has been already shown, both experimentally<sup>30</sup> and theoretically.<sup>19,28,35–37</sup> Our study reveals that the description of the zinc binding site depends upon the choice of the metal ligands environment included in the model, and more precisely around the Zn–His119–Glu117 triad, which were not included in previous studies. It seems that an unquestionable model should include all groups bound to the first shell, especially those including charge and those surrounding charge, even if the minimum required is still unclear. Indeed, modeling the surrounding protein by an isotropic polarizable continuum seems to be inadequate as it cannot replace reliably the effect of all the local interactions due to the hydrogen bond network.

Our models **1–10** do not include the peptidic main chain which plays a major role in the rigidity of the enzyme and could influence its reactivity.<sup>38–40</sup> As a consequence, these models are very flexible and lead to optimized structures in which ligand orientation for metal coordination is not exactly equivalent to the experimental orientation. This lack of rigidity is also illustrated by the disruption of the hydrogen bond between the Zn-bound hydroxy and the MeOH, model of the side chain of Thr199, in **4–5.OH<sup>−</sup>** and **8–10.OH<sup>−</sup>** (thus not included in this study) although this interaction is supposed to be present in the enzyme. As noted previously, one role of the indirect ligands has been proposed to be their capability to orient first-shell ligands for suitable metal coordination for catalysis.

Comparison between models **10** and **11** permit to evaluate the effect of the peptidic constraint on the relative stability of the OH and NH isomers. The constraint modifies the interaction between the chemical groups of the Zn–His–Glu triad, preventing the carboxylate moiety of the NH isomer from maximizing its interaction with the imidazole ring, but leaving it more accessible to the solvent. As a consequence, the rigidity of the system strongly favors the OH isomer in the gas phase, but increases its solvation energy. Overall, the effect of this orientation constraint on the relative energy of the OH and NH isomers appears to be small.

## 2. Relevance of proton transfer to the HCA II catalytic cycle

The present results shed new light on the nature of the HCA II active site and its activation mechanism. The active hydroxy form of the enzyme (**CA.OH<sup>−</sup>**) presents, as usually considered, a zinc metal cation surrounded by three imidazole ligands and an OH<sup>−</sup> group. The rest aqua form (**CA.H<sub>2</sub>O**) seems however to be able to adopt various structures. Indeed, the OH and the NH isomers, based on our calculations, have close thermodynamic stability and thus both of them should be able to correspond to the rest form of the enzyme, even if a preference for the OH isomer emerges from our larger model. We suggest that the enzyme takes advantage of this possibility of fine tuning the charge transfer of the His119–Glu117 dyad to the zinc cation in its very efficient catalytic cycle.

From the zinc-aqua resting state of the enzyme, the first step of the hydration of carbon dioxide is the deprotonation of the Zn-bound water molecule to form the zinc-hydroxy state. This rate-limiting proton transfer from metal-bound H<sub>2</sub>O to the solvent is assumed to determine the observed pK<sub>a</sub> of the catalytic site of about 7. As the zinc-bound water pK<sub>a</sub> reflects the solvated free energy difference between the zinc-aqua and zinc-hydroxy forms, our results suggest that this could be determined by the free energy difference between the **CA.OH<sup>−</sup>.H<sub>2</sub>O** and **CA.NH<sup>−</sup>.OH<sup>−</sup>** structures, indicating a proton shift between His119 and Glu117 concomitant with the proton transfer from the zinc-bound water to His64. This proton shift would lead to a decrease of the charge transfer for the His119–Glu117 dyad to the metal, thus increasing its Lewis acidity required to stabilize the hydroxide group. The direct and indirect ligands of the Zn<sup>2+</sup> ion have biological importance as demonstrated by site-specific mutagenesis on carbonic anhydrases.<sup>30,41–43</sup> The interpretation of experimental results on their influence on the zinc-bound water pK<sub>a</sub> remains unsatisfactory, however.<sup>44</sup> The His96Cys and His119Asp CAII mutants have a negatively-charged group replacing a neutral histidine ligand, leading to a decrease of the ability of zinc to polarize its bound water molecule and thus to an increase of the zinc-bound water pK<sub>a</sub> (8.5 and 8.6, respectively, compared to the wild-type pK<sub>a</sub> of 6.8).<sup>45</sup> To a lower extent, the same electronic effect is observed for indirect ligands mutagenesis. Indeed, the Gln92Leu and Gln92Glu CAII mutants have a zinc-bound water pK<sub>a</sub> of 6.4 and 7.7, respectively,<sup>41</sup> reflecting respectively the disruption or the strengthening of the H-bond between His94 and Gln92 and thus the fine tuning of the His–Zn interaction. The Glu117Gln CAII mutant presents a pK<sub>a</sub> value >9.9 which has been explained by a reversed polarity of the His119–Gln117 hydrogen bond.<sup>46</sup> However, this did not explain the greater effect of glutamate–histidine (Glu117<sup>(−)</sup>–His119<sup>(0)</sup>) to glutamine–histidinate (Gln117<sup>(0)</sup>–His119<sup>(−)</sup>) dyad shift compared to histidine to cysteinate (His<sup>(0)</sup>96Cys<sup>(−)</sup>) or aspartate (His<sup>(0)</sup>119Asp<sup>(−)</sup>) shift. Indeed, going from neutral to anionic zinc-ligand on both the zinc-aqua and zinc-hydroxy form should only increase the pK<sub>a</sub> by less than 2 units as in the His96Cys and His119Asp CAII mutants. Our results suggest that Glu117Gln-CA is characterized, compared to the wild-type enzyme, by the neutral to anionic zinc-ligand shift *only* for the zinc-hydroxy form as the wild-type zinc-aqua form is already able to possess a Zn<sup>2+</sup>–Im<sup>−</sup>⋯MeCOOH<sup>0</sup> unit. Our computations assess the energy difference between **CA.NH<sup>−</sup>.OH<sup>−</sup>** and **CA.OH<sup>−</sup>.OH<sup>−</sup>** to be about 15–20 kJ mol<sup>−1</sup>, corresponding to an increase of about three pK<sub>a</sub> units for the hypothetical **CA.OH<sup>−</sup>.H<sub>2</sub>O/CA.OH<sup>−</sup>.OH<sup>−</sup>** acid/base couple compared to the natural **CA.OH<sup>−</sup>.H<sub>2</sub>O/CA.NH<sup>−</sup>.OH<sup>−</sup>** one. This result agrees well with the experimental pK<sub>a</sub> value >9.9 of the Glu117Gln CAII mutant corresponding to the **Glu117Gln-CA.OH<sup>−</sup>.H<sub>2</sub>O/Glu117Gln-CA.OH<sup>−</sup>.OH<sup>−</sup>** acid/base couple.

In the next step of the catalytic cycle, the zinc-bound hydroxide attacks the carbon atom of CO<sub>2</sub> to form zinc-bound bicarbonate. Contrary to previous studies on [Zn(NH<sub>3</sub>)<sub>3</sub>(OH)]<sup>+</sup> models which show quite large barrier values,<sup>47–49</sup> recent theoretical studies indicate that the use of imidazole ligands instead of NH<sub>3</sub> noticeably increases the nucleophilicity of the Zn-bound



OH<sup>−</sup> group and thus decreases the activation barrier.<sup>28,50</sup> Furthermore, the inclusion of protein environment in QM/MM calculation leads, depending on the backbone conformation, to a barrierless diffusion and addition of the carbon dioxide to the zinc-bound oxygen.<sup>39</sup> This indicates that the Zn-bound hydroxide does not require an increased nucleophilicity to complete hydration of CO<sub>2</sub>. This is consistent with our results which indicate a greater stabilization of the NH compared to the OH isomer for the Zn-bound hydroxide form. On the other hand, it has been recently proposed an alternative catalytic mechanism in which the nucleophilic attack on carbon dioxide is provided by zinc-bound water.<sup>51</sup> In order to make the coordinated water sufficiently nucleophilic to attack CO<sub>2</sub>, the enzyme may play a role in dispersing the positive charge of Zn<sup>2+</sup>–OH<sub>2</sub> through a hydrogen bond network. This could be achieved through the Thr199–Glu106 dyad, hydrogen bound to the Zn-bound water, but also by the His119–Glu106 one as shown by our calculations. Indeed, the zinc-bound water nucleophilicity would be greater in the OH isomer due to decrease of the Zn<sup>2+</sup> Lewis acidity, a noticeable part of the positive charge being delocalized to its anionic histidinate ligand.<sup>52</sup>

The return to the zinc-water resting state is achieved by the rearrangement of the linked bicarbonate followed by its displacement by a water molecule.<sup>28,49,53,54</sup> While there is no X-ray structure of HCA II with HCO<sub>3</sub><sup>−</sup> bound to zinc, crystallographic studies of a Thr200His mutant of HCA II<sup>55</sup> and cobalt substituted HCA II<sup>56</sup> complexed with HCO<sub>3</sub><sup>−</sup> indicate a bidentate or pseudo-bidentate bicarbonate ligand to the metal. This high affinity of bicarbonate for metal binding is not in favour of an easy substitution of the anionic HCO<sub>3</sub><sup>−</sup> by a neutral H<sub>2</sub>O from the Lewis acid Zn<sup>2+</sup>. This displacement could be facilitated by a bicarbonate shift from bidentate to monodentate coordination thus leaving an open site for the entering water molecule. We suggest that this fine-tuning coordination mode could be correlated with the proton transfer from Glu117 to His119. Indeed, from a Zn<sup>2+</sup>–ImH<sup>0</sup>...MeCOO<sup>−</sup> unit, the shuttling of the proton between the H-bonded atoms of the Glu117–His119 couple will decrease the zinc charge and thus its affinity to bicarbonate which will concomitantly shift to a monodentate coordination to zinc. Our hypothesis is supported by the experimental observation on bicarbonate binding to the non-heme iron of photosystem II which shows that Fe<sup>2+</sup> oxidation to Fe<sup>3+</sup> is accompanied by both bidentate to monodentate coordination shift of the bicarbonate and deprotonation of one histidine ligand.<sup>12</sup> This assistance to ligand coordination change, by proton transfer in a Zn–His–Glu/Asp triad, may bear some generality. It may be postulated to occur, for instance, in the carboxylate shift.<sup>57,58</sup> In the active site of farnesyltransferase, a Zn–His–Asp triad is present and may assist the bidentate to monodentate switch of another Asp residue bound to zinc.

The displacement of bicarbonate by water, facilitated by the adaptability of the Zn–His119–Glu117 triad, leads back to the Zn-aqua resting state of the enzyme in its OH isomer as previously described.

## Conclusion

From reliable models of the active site of HCA II, we have shown that the position of the proton in the His119–Glu117

dyad is greatly dependent upon both the model used and the nature of the zinc-bound oxygen ligand. As a consequence, an extended model for the HCA II zinc site should include all groups bound to the first shell, especially those including charge and those surrounding charge, in order to obtain accurate information. Furthermore, our results indicate that the isomerisation between the Zn<sup>2+</sup>–ImH<sup>0</sup>...MeCOO<sup>−</sup> and the Zn<sup>2+</sup>–Im<sup>−</sup>...MeCOOH<sup>0</sup> forms could be used by the enzyme along the catalytic cycle to facilitate the deprotonation of the Zn-bound water (O to N proton transfer) and the displacement of the bicarbonate by water (N to O proton transfer). His119–Glu117 could thus be described not only as a spectator zinc-ligand which does not participate in the reaction mechanism, but also as a tunable ligand which could take two different forms during the catalytic cycle. Quantum molecular dynamics simulations could, in principle, provide a deeper insight into this mechanism. However, such calculations on extended model like **12** remain currently out of reach. Such expensive calculations are now available for small to medium-size models, but we have shown that such models may be inadequate to properly describe the anisotropy of the protein environment. In any case, the basic picture of His119–Glu117 tunability appears to be well established by the present calculations.

## Methods

### 1. QM calculations

Two basis sets were used in this study. For geometry optimisations and vibrational frequency calculations, the 6-31G\* basis was used for H, C, N, and O, and the Wachters [14s9p5d1f/9s5p3d1f] basis was used for Zn.<sup>59</sup> This is referred to as BS1. For final energy calculations, basis set 2 (noted BS2) consists in the 6-311 + G(2d,2p) for H, C, N and O, and the extended Wachters basis [15s11p6d2f/10s7p4d2f] for Zn.

Full geometry optimisations without symmetry constraints were carried out with the Gaussian 03 program suite<sup>60</sup> at the Hartree–Fock HF/BS1 level for complexes **1–10**. Frequency analyses were carried out to confirm that the reported structures are minima or transition states on the HF/BS1 potential energy surface, and to evaluate the thermal and entropic contributions necessary for the calculation of enthalpies and free energies. The positions of eight carbon atoms were frozen for model **11** as in the X-ray structure during the geometry optimisation process. Enthalpies and gas phase free energies for **11** have been derived from thermal and entropic contributions computed for **10**. Final energies for models **1–11** were obtained at the density functional B3LYP level and, in some cases, post-HF MP2 levels at the HF/BS1 geometries, levels denoted as B3LYP/BS2//HF/BS1 and MP2/BS2//HF/BS1 respectively.

These levels of calculation have been shown to give reasonable geometries and accurate final energies.<sup>33</sup> This is confirmed for the energetic data examined in this study by geometry optimisation of complexes **2** and **11** at the B3LYP/BS1 level (see Table 1).

For several cases (**1**·H<sub>2</sub>O, **2**·H<sub>2</sub>O and **4**·H<sub>2</sub>O), the NH isomer could not be located as a minimum, the optimization

process leading to the OH isomer. In order to check the reliability of the HF/BS1 methodology in these cases, geometry optimizations were carried out at the more accurate MPWB1K<sup>61</sup> and MP2/6-31+G(d,p) (Wachters [15s11p6d1f/10s7p4d1f] basis for Zn) levels. The basis set is expanded by adding p polarization functions on hydrogens since isomerization between NH and OH isomers is a proton transfer, and adding diffuse functions on all heavy atoms since negative charge re-location accompanies proton transfer. It turned out to be impossible to locate a minimum for the NH isomer in all cases due to proton transfer from imidazole to carboxylate, thus confirming the results obtained at the HF/BS1 level.

Due to the flexibility of models **1–10**, great care has been taken to ensure that interactions between successive shells be preserved in all cases. Optimized geometries of the OH and NH isomers of each complex have been compared to ascertain their structural equivalence (see Table S1 in the ESI†) and thus the reliability of the relative energy comparison. Furthermore, structural comparison has been done between optimized and X-ray crystallographic structures in order to check the preservation of environment structure. Whenever a large discrepancy was found, the model was excluded from the discussion (*vide supra*). Fixing the positions of several carbon atoms in the periphery of model **11** ensures a geometry that remains close to the X-ray structure.

Solvation effects were included with conductor polarized continuum method (CPCM)<sup>62,63</sup> calculations as implemented in the Gaussian 03 software. In this method the solute cavities are modelled on the optimized molecular shape, and include both electrostatic and non-electrostatic contributions to the energies. The CPCM calculations were performed on the HF/BS1 geometries at the HF level with BS1+ derived from BS1 by addition of one set of diffuse functions on all atoms except H. In the CPCM calculations the area of the tesserae is set at 0.2 Å<sup>2</sup>, the solvent is water ( $\epsilon = 78.39$ ) and the molecular cavity is built up by the UAKS model.<sup>64</sup> The validity of this solvation approach has been demonstrated before in the study of zinc complexes.<sup>65</sup> CPCM calculation have also be done with a dielectric constant equal to 4 to model the effect of a protein environment on a buried Zn active site.<sup>66–68</sup>

## 2. QM/QM' calculations

In order to provide a more accurate description of the environment of the enzyme, we carried out QM/QM' calculations on a large fraction of the enzyme using a two-layer ONIOM procedure.<sup>69</sup> In this approach, the system under study, labelled **12**, is split into two subsystems: The QM system (inner layer, 101 atoms including 15 H junction atoms) corresponds to model **11** and is computed at the HF/BS1 level. The QM' system (outer layer, further 876 atoms) consists of residues and water molecules possessing at least one atom within 10 Å of the zinc atom and is computed at the PM3 level including the optional molecular mechanics correction for HCON linkages (noted PM3MM).<sup>70–72</sup> This level has been previously used successfully in the study of the farnesyl transferase zinc enzyme.<sup>73</sup> The geometry of model **12** was optimised by fixing the position of all C<sub>α</sub> atoms of the outer

layer (59 atoms) in their X-ray structure position. This permits to keep the enzymatic environment of the active site in place without introducing any bias in the optimisation process of the inner layer. Final energies for the inner shell of model **12** were obtained at the B3LYP/BS2 level, as well as, for the entire system, at the Hartree–Fock and at the density functional B3LYP level including the 977 atoms of the system. For this HF calculation, BS2 and 3-21G were used for atoms belonging respectively to the inner and outer layers, thus leading to a total of 6546 AOs. For the B3LYP calculation, STO-3G was used instead of 3-21G for the atoms of the outer layer, reducing to 4341 the number of AOs. Enthalpies and gas phase free energies for **12** have been derived from thermal and entropic contributions computed for **10**. Solvation effects were included with the CPCM method computed at the PM3MM level for the entire model **12**.

## Acknowledgements

We are grateful to the Centre National de la Recherche Scientifique (CNRS) and Ecole Polytechnique for financial support. We also thank the Institut du Développement et des Ressources en Informatique Scientifique (IDRIS, grant number 0543) and the Centre Informatique National de l'Enseignement Supérieur (CINES, grant number dcm2335) for computational resources.

## References

- W. N. Lipscomb and N. Strater, *Chem. Rev.*, 1996, **96**, 2375–2433.
- W. R. Chegwidden, N. D. Carter and Y. H. Edwards, *The Carbonic Anhydrases: New Horizons*, Birkhäuser Verlag, Basel, Switzerland, 2000.
- S. Lindskog, *Pharmacol. Ther.*, 1997, **74**, 1–20.
- K. Hakansson, M. Carlsson, L. A. Svensson and A. Liljas, *J. Mol. Biol.*, 1992, **227**, 1192–1204.
- H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
- E. Meyer, *Protein Sci.*, 1992, **1**, 1543–1562.
- P. Hudaky and A. Perczel, *J. Phys. Chem. A*, 2004, **108**, 6195–6205.
- K. Hasegawa, T.-A. Ono and T. Noguchi, *J. Phys. Chem. A*, 2002, **106**, 3377–3390.
- Y. Bu and R. I. Cukier, *J. Phys. Chem. B*, 2004, **108**, 10089–10100.
- J. S. de Ropp, S. Sham, A. Asokan, S. Newmyer, P. R. Ortiz de Montellano and G. N. La Mar, *J. Am. Chem. Soc.*, 2002, **124**, 11029–11037.
- G. Smulevich, J. M. Mauro, L. A. Fishel, A. M. English, J. Kraut and T. G. Spiro, *Biochemistry*, 1988, **27**, 5477–5485.
- R. Hienerwadel and C. Berthomieu, *Biochemistry*, 1995, **34**, 16288–16297.
- C. Berthomieu and R. Hienerwadel, *Biochemistry*, 2001, **40**, 4044–4052.
- M. Iwaki, G. Yakovlev, J. Hirst, A. Osyczka, P. L. Dutton, D. Marshall and P. R. Rich, *Biochemistry*, 2005, **44**, 4230–4237.
- L. Marboutin, A. Boussac and C. Berthomieu, *JBIC, J. Biol. Inorg. Chem.*, 2006, **11**, 811–823.
- S. Franzen, *J. Am. Chem. Soc.*, 2001, **123**, 12578–12589.
- K. P. Jensen and U. Ryde, *Mol. Phys.*, 2003, **101**, 2003–2018.
- L. Capece, M. A. Marti, A. Crespo, F. Doctorovich and D. A. Estrin, *J. Am. Chem. Soc.*, 2006, **128**, 12455–12461.
- M. Krauss and D. R. Garmer, *J. Am. Chem. Soc.*, 1991, **113**, 6426–6435.
- J. El Yazal and Y. P. Pang, *J. Phys. Chem. B*, 1999, **103**, 8773–8779.

- 21 J. El Yazal, R. R. Roe and Y. P. Pang, *J. Phys. Chem. B*, 2000, **104**, 6662–6667.
- 22 T. Dudev and C. Lim, *J. Phys. Chem. B*, 2001, **105**, 4446–4452.
- 23 T. Dudev and C. Lim, *J. Chin. Chem. Soc.*, 2003, **50**, 1093–1102.
- 24 Y. Lin and C. Lim, *J. Am. Chem. Soc.*, 2004, **126**, 2602–2612.
- 25 Y.-L. Lin, Y.-M. Lee and C. Lim, *J. Am. Chem. Soc.*, 2005, **127**, 11336–11347.
- 26 F. L. Gervasio, V. Schettino, S. Mangani, M. Krack, P. Carloni and M. Parrinello, *J. Phys. Chem. B*, 2003, **107**, 6886–6892.
- 27 T. Stams and D. W. Christianson, in *The Carbonic Anhydrases: New Horizons*, ed. W. R. Chegwidden, N. D. Carter and Y. H. Edwards, Birkhäuser Verlag, Basel, Switzerland, 2000, pp. 159–174.
- 28 A. Bottoni, C. Z. Lanza, G. P. Miscione and D. Spinelli, *J. Am. Chem. Soc.*, 2004, **126**, 1542–1550.
- 29 P. Argos, R. M. Garavito, W. Eventoff, R. M. G. and C. I. Brändén, *J. Mol. Biol.*, 1978, **126**, 141–158.
- 30 D. W. Christianson and C. A. Fierke, *Acc. Chem. Res.*, 1996, **29**, 331–339.
- 31 T. Dudev, Y.-L. Lin, M. Dudev and C. Lim, *J. Am. Chem. Soc.*, 2003, **125**, 3168–3180.
- 32 D. W. Christianson and R. S. Alexander, *J. Am. Chem. Soc.*, 1989, **111**, 6412–6419.
- 33 G. Frison and G. Ohanessian, *J. Comput. Chem.*, 2008, **29**, 416–433.
- 34 M. Kaukonen, P. Söderhjelm, J. Heimdal and U. Ryde, *J. Chem. Theory Comput.*, 2008, **4**, 985–1001.
- 35 S. Alvarez-Santos, A. Gonzalez-Lafont and J. M. Lluch, *Can. J. Chem.*, 1998, **76**, 1027–1032.
- 36 T. Marino, N. Russo and M. Toscano, *J. Am. Chem. Soc.*, 2005, **127**, 4242–4253.
- 37 G. P. Miscione, M. Stenta, D. Spinelli, E. Anders and A. Bottoni, *Theor. Chem. Acc.*, 2007, **118**, 193–201.
- 38 P. Comba, A. Lledos, F. Maseras and R. Remenyi, *Inorg. Chim. Acta*, 2001, **324**, 21–26.
- 39 M. J. Loferer, C. S. Tautermann, H. H. Loeffler and K. R. Liedl, *J. Am. Chem. Soc.*, 2003, **125**, 8921–8927.
- 40 R. J. F. Branco, P. A. Fernandes and M. J. Ramos, *THEOCHEM*, 2005, **729**, 141–146.
- 41 L. L. Kiefer, S. A. Paterno and C. A. Fierke, *J. Am. Chem. Soc.*, 1995, **117**, 6831–6837.
- 42 C. A. Lesburg and D. W. Christianson, *J. Am. Chem. Soc.*, 1995, **117**, 6838–6844.
- 43 A. K. Mohanty, K. K. Kannan and S. K. Mahajan, *J. Biosci.*, 1998, **23**, 235–246.
- 44 D. Lu and G. A. Voth, *Proteins: Struct., Funct., Genet.*, 1998, **33**, 119–134.
- 45 L. L. Kiefer and C. A. Fierke, *Biochemistry*, 1994, **33**, 15233–15240.
- 46 C.-C. Huang, C. A. Lesburg, L. L. Kiefer, C. A. Fierke and D. W. Christianson, *Biochemistry*, 1996, **35**, 3439–3446.
- 47 Y.-J. Zheng and K. M. Merz, Jr, *J. Am. Chem. Soc.*, 1992, **114**, 10498–10507.
- 48 C. Muguruma, *THEOCHEM*, 1999, **461–462**, 439–452.
- 49 M. Mauksch, M. Brauer, J. Weston and E. Anders, *ChemBioChem*, 2001, **2**, 190–198.
- 50 M. Bräuer, J. L. Pérez-Lustres, J. Weston and E. Anders, *Inorg. Chem.*, 2002, **41**, 1454–1463.
- 51 S. Thoms, *J. theor. Biol.*, 2002, **215**, 399–404.
- 52 X. Zhang, C. D. Hubbard and R. van Eldik, *J. Phys. Chem.*, 1996, **100**, 9161–9171.
- 53 K. M. Merz, Jr and L. Banci, *J. Am. Chem. Soc.*, 1997, **119**, 863–871.
- 54 C. S. Tautermann, M. J. Loferer, A. F. Voegelé and K. R. Liedl, *J. Phys. Chem. B*, 2003, **107**, 12013–12020.
- 55 Y. Xue, J. Vidgren, L. A. Svensson, A. Liljas, B.-H. Jonsson and S. Lindskog, *Proteins: Struct., Funct., Genet.*, 1993, **15**, 80–87.
- 56 K. Hakansson and A. Wehnert, *J. Mol. Biol.*, 1992, **228**, 1212–1218.
- 57 T. Dudev and C. Lim, *Acc. Chem. Res.*, 2007, **40**, 85–93.
- 58 S. F. Sousa, P. A. Fernandes and M. J. Ramos, *J. Am. Chem. Soc.*, 2007, **129**, 1378–1385.
- 59 A. J. Wachtters, *J. Chem. Phys.*, 1970, **52**, 1033–1036.
- 60 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03 (Revision B.05)*, Gaussian, Inc, Pittsburgh, PA, 2003.
- 61 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 6908–6918.
- 62 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
- 63 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.
- 64 Y. Takano and K. N. Houk, *J. Chem. Theory Comput.*, 2005, **1**, 70–77.
- 65 D. Picot, G. Ohanessian and G. Frison, *Inorg. Chem.*, 2008, **47**, 8167–8178.
- 66 S. C. Harvey and P. Hoekstra, *J. Phys. Chem.*, 1972, **76**, 2987–2994.
- 67 M. K. Gilson and B. H. Honig, *Biopolymers*, 1986, **25**, 2097–2119.
- 68 A. Warshel, P. K. Sharma, M. Kato and W. W. Parson, *Biochim. Biophys. Acta*, 2006, **1764**, 1647–1676.
- 69 F. Maseras and K. Morokuma, *J. Comput. Chem.*, 1995, **16**, 1170–1179.
- 70 J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 209–220.
- 71 J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 221–264.
- 72 J. J. P. Stewart, *J. Comput. Chem.*, 1991, **12**, 320–341.
- 73 S. F. Sousa, P. A. Fernandes and M. J. Ramos, *J. Comput. Chem.*, 2007, **28**, 1160–1168.